Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Linear Algebra and its Applications 364 (2003) 13–31

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Characterizations and lower bounds for the spread of a normal matrix

Jorma Kaarlo Merikoski [a,*], Ravinder Kumar [b]

[a]*Department of Mathematics, Statistics and Philosophy, University of Tampere, Tampere FIN-33014, Finland*

[b]*Department of Mathematics, Dayalbagh Educational Institute, Dayalbagh, Agra 282005, Uttar Pradesh, India*

Received 18 February 2002; accepted 26 July 2002

Submitted by G.P.H. Styan

Abstract

The spread of an $n \times n$ matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ is defined by $\operatorname{spr}\mathbf{A} = \max_{j,k} |\lambda_j - \lambda_k|$. We prove that if $\mathbf{A}$ is normal, then

$$\begin{aligned}
\operatorname{spr}\mathbf{A} &= \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathbf{y}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1\right\} \\
&= \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \operatorname{re}\mathbf{x}^*\mathbf{y} = 0\right\} \\
&= \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\} \\
&= \max\left\{\operatorname{spr}\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2} \mid z \in \mathbf{C}, |z| = 1\right\} \\
&= \sqrt{2}\max\left\{\left(|\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2| + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x} - |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2\right)^{1/2} \mid \mathbf{x} \in \mathbf{C}^n, \|\mathbf{x}\| = 1\right\}.
\end{aligned}$$

We also present several lower bounds for $\operatorname{spr}\mathbf{A}$, given by these characterizations.

Keywords: Normal matrices; Eigenvalues; Spread

* Corresponding author.

E-mail addresses: jorma.merikoski@uta.fi (J.K. Merikoski), drravinderkumar@yahoo.com (R. Kumar).

0024-3795/03/$ - see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0024-3795(02)00534-7

1. Notations and preliminaries

We denote complex vectors by bold lower case letters and complex matrices by bold upper case letters. Identifying n-dimensional vectors with $n \times 1$ matrices, we denote the Euclidean inner product of vectors $\mathbf{x}$ and $\mathbf{y}$ by $\mathbf{x}^*\mathbf{y}$ and the Euclidean norm of $\mathbf{x}$ by $\|\mathbf{x}\| = \sqrt{\mathbf{x}^*\mathbf{x}}$. Furthermore, $\|\mathbf{x}\|_1$ is the l_1-norm of $\mathbf{x}$, $|S|$ is the number of elements of a finite set S, and $[n] = \{1, \ldots, n\}$. If S is a set of complex numbers, re S is the set of real parts of its elements, $\operatorname{diam} S = \sup\{|u - v| \mid u, v \in S\}$ is its diameter, and co S is its convex hull. The conjugates $\bar{\mathbf{x}}$ and $\bar{\mathbf{A}}$ are understood elementwise.

The spread of an $n \times n$ matrix $\mathbf{A}$ ($n \geqslant 2$) with spectrum $\operatorname{spec} \mathbf{A} = \{\lambda_1, \ldots, \lambda_n\}$ is defined by

$$\operatorname{spr} \mathbf{A} = \max_{j,k} |\lambda_j - \lambda_k|.$$

If $\operatorname{spec} \mathbf{A}$ is real, we order $\lambda_1 \geqslant \cdots \geqslant \lambda_n$, and so

$$\operatorname{spr} \mathbf{A} = \lambda_1 - \lambda_n.$$

We denote by $\mathbf{u}_j$ a unit eigenvector of $\mathbf{A}$ corresponding to λ_j.

The vector $\mathbf{e}_j$ is the jth standard basis vector ($1 \leqslant j \leqslant n$). If $(\emptyset \neq) J \subset [n]$, we denote $\mathbf{e}_J = \sum_{j \in J} \mathbf{e}_j$ and $\mathbf{e} = \mathbf{e}_{[n]} = (1, \ldots, 1)^{\mathrm{T}}$. If $\mathbf{B}$ is an $n \times m$ matrix, then su $\mathbf{B}$ denotes the sum of its elements and $\mathbf{B}_{JK}$ denotes its submatrix with rows in $J \subset [n]$ and columns in $K \subset [m]$ ($J, K \neq \emptyset$), and $\mathbf{B}_K = \mathbf{B}_{KK}$.

We need some simple properties of the numerical range

$$F(\mathbf{A}) = \left\{\mathbf{x}^*\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbf{C}^n, \|\mathbf{x}\| = 1\right\}$$

of an $n \times n$ matrix $\mathbf{A}$.

Proposition 1 (e.g., [7, Property 1.2.5]). *If* $\mathbf{A}$ *is a square matrix, then*

$$F\left(\frac{\mathbf{A} + \mathbf{A}^*}{2}\right) = \operatorname{re} F(\mathbf{A}).$$

Proposition 2 (e.g., [7, Theorem 1.2.9 and p. 12]). *If* $\mathbf{A}$ *is a square matrix, then* $F(\mathbf{A}) \supset \operatorname{co}(\operatorname{spec} \mathbf{A})$ *and so* $\operatorname{spr} \mathbf{A} \leqslant \operatorname{diam} F(\mathbf{A})$. *If* $\mathbf{A}$ *is normal, then* $F(\mathbf{A}) = \operatorname{co}(\operatorname{spec} \mathbf{A})$ *and so* $\operatorname{spr} \mathbf{A} = \operatorname{diam} F(\mathbf{A})$.

Proposition 3 (e.g., [7, Theorem 1.2.11]). *If* $\mathbf{B}$ *is a principal submatrix of a square matrix* $\mathbf{A}$, *then* $F(\mathbf{B}) \subset F(\mathbf{A})$.

The last statement of Proposition 2 (not included in [7]) relies on the fact that, given complex numbers $z_1, \ldots, z_m$, the set $H = \operatorname{co}\{z_1, \ldots, z_m\}$ has $\operatorname{diam} H = \max\{|z_j - z_k| \mid 1 \leqslant j, k \leqslant m\}$. This geometrically obvious fact is crucial to us, and therefore we prove it algebraically.

Standard continuity and compactness arguments imply that $\operatorname{diam} H = \max\{|u - v| \mid u, v \in H\}$. We show by contradiction that the maximizing u and v are extreme points of H. Assume that v is not an extreme point. Then there exist $v_1, v_2 \in H$, such that v_1, v_2, and u are all unequal and that $v = (v_1 + v_2)/2$. Now

$$\begin{aligned}|u - v| = \tfrac{1}{2}|u - v_1 + u - v_2| &\leqslant \tfrac{1}{2}(|u - v_1| + |u - v_2|)\\ &\leqslant \max(|u - v_1|, |u - v_2|).\end{aligned}$$

The first inequality is strict if $\arg(u - v_1) \neq \arg(u - v_2)$, and the second inequality is strict otherwise. Thus $|u - v|$ can be increased, and so it is not the maximum.

Proposition 2 and 3 imply

Proposition 4. *If* $\mathbf{B}$ *is a principal submatrix of a normal matrix* $\mathbf{A}$, *then* $\operatorname{spr}\mathbf{B} \leqslant \operatorname{spr}\mathbf{A}$.

This is not valid for all square matrices. For a counterexample, let

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $\operatorname{spec}\mathbf{A} = \{0, 0, 0\}$ and $\operatorname{spec}\mathbf{B} = \{1, -1\}$. Hence $\operatorname{spr}\mathbf{A} = 0$ but $\operatorname{spr}\mathbf{B} = 2$.

We state also the following elementary proposition.

Proposition 5. *If* w *is a complex number*, *there exists a complex number* z *satisfying* $|z| = 1$ *and* $zw = |w|$.

Proof. If $w \neq 0$, then $z = \bar{w}/|w|$; if $w = 0$, then any z with $|z| = 1$ works. $\square$

2. Introduction

There are several bounds for the spread in the literature, see the reference list. It is interesting that the spread of an $n \times n$ matrix is the spectral radius of an $n^2 \times n^2$ matrix [7, Problem 4.3.7]. Thus all that is known about estimating the spectral radius can, in principle, be applied in estimating the spread, but in practice this approach is too complicated. If all the eigenvalues are real, then upper and lower bounds for the largest and smallest eigenvalue give bounds for the spread as corollaries, see e.g., [20, Theorem 2.5].

Trivially [6, Problem 4.2.5], for $1 \leqslant j \leqslant n$,

$$\min\left\{|\mathbf{x}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x} \in \mathbf{C}^n, \|\mathbf{x}\| = 1\right\} \leqslant |\lambda_j| \leqslant \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x} \in \mathbf{C}^n, \|\mathbf{x}\| = 1\right\}.$$

If $\mathbf{A}$ is Hermitian, then (see e.g., [6, Theorem 4.2.2])

$$\lambda_1 = \max\left\{\mathbf{x}^*\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbf{C}^n, \|\mathbf{x}\| = 1\right\}, \quad \lambda_n = \min\left\{\mathbf{x}^*\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbf{C}^n, \|\mathbf{x}\| = 1\right\}.$$

Thus we expect that the difference $\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathbf{y}^*\mathbf{A}\mathbf{y}$ has a role in characterizing $\operatorname{spr}\mathbf{A}$. Indeed, Mirsky [16, p. 598, Corollary] proved for normal matrices that

$$(\mathrm{C}_1) \quad \operatorname{spr}\mathbf{A} = \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathbf{y}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}.$$

(Actually he used sup instead of max, but we prefer here and in related discussions max, which can be easily shown to exist.) Particular choices of $\mathbf{x}$ and $\mathbf{y}$ give lower bounds for $\operatorname{spr}\mathbf{A}$. We will see (Theorem 1) that the condition $\mathbf{x}^*\mathbf{y} = 0$ can be dropped out, and so we will get more freedom in choosing $\mathbf{x}$ and $\mathbf{y}$.

Mirsky [16, Lemma 1] proved for normal matrices also that

$$(\mathrm{C}_2) \quad \operatorname{spr}\mathbf{A} \geqslant \max\left\{\operatorname{spr}\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2} \mid z \in \mathbf{C}, |z| = 1\right\}.$$

If we know a lower bound for the spread of a Hermitian matrix, we can apply this inequality to find lower bounds for the spread of a normal matrix. We will see (Theorem 2) that in fact equality holds in (C_2).

Furthermore, Mirsky [16, Theorem 1] proved for Hermitian matrices that

$$(\mathrm{C}_3) \quad \operatorname{spr}\mathbf{A} = 2\max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}$$

and [16, (4)] for normal matrices that

$$(\mathrm{C}_3') \quad \operatorname{spr}\mathbf{A} \geqslant \sqrt{3}\max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}.$$

We will (Theorem 3) generalize (C_3) for normal matrices. Mirsky [16, p. 593] showed that $\sqrt{3}$ is the best possible coefficient in (C_3') in the sense that equality is attained for some $\mathbf{A}$ ($n \geqslant 3$) with $\operatorname{spr}\mathbf{A} > 0$. Since also strict inequality is attained in (C_3') by (C_3) if $\mathbf{A}$ is Hermitian with $\operatorname{spr}\mathbf{A} > 0$, the inequality ($\mathrm{C}_3'$) neither characterizes $\operatorname{spr}\mathbf{A}$ nor can be sharpened to an equality with this property.

Bloomfield and Watson [3, (5.3)] proved for real symmetric matrices that

$$(\mathrm{C}_4) \quad \operatorname{spr}\mathbf{A} \geqslant 2\max\left\{\left(\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2\right)^{1/2} \mid \mathbf{x} \in \mathbf{R}^n, \|\mathbf{x}\| = 1\right\},$$

rediscovered by Styan [18, Theorem 1]; see also [8, Section 5.4; 9]. We will note (Theorem 4) that also (C_4) is actually equality and that it is valid for all Hermitian matrices (then $\mathbf{x} \in \mathbf{C}^n$). We will also generalize (Corollary 4) it for normal matrices. Particular choices of $\mathbf{x}$ give lower bounds for $\operatorname{spr}\mathbf{A}$.

We will first (Sections 3–5) present Theorems 1–4 and Corollary 4 outlined above. Second (Sections 6–10), we will apply these theorems to find lower bounds for the spread. Thereafter (Section 11) we will study two other well-known bounds from our point of view. Next (Section 12), we will note that some of our bounds involving element sums have analogies involving traces. Finally, after some remarks (Section 13), we will (Section 14) compare different bounds empirically.

3. Modifying (C_1) and improving (C_2)

Theorem 1. *If $\mathbf{A}$ is a normal $n \times n$ matrix, then*

$$\operatorname{spr}\mathbf{A} = \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathbf{y}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1\right\}.$$

Proof. An equivalent claim,

$$\operatorname{spr}\mathbf{A} = \max\left\{|u - v| \mid u, v \in F(\mathbf{A})\right\},$$

is true by Proposition 2. □

Theorem 2. *If* $\mathbf{A}$ *is a normal matrix, then*

$$\operatorname{spr}\mathbf{A} = \max\left\{\operatorname{spr}\frac{z\,\mathbf{A} + \bar{z}\mathbf{A}^*}{2} \mid z \in \mathbf{C}, |z| = 1\right\}.$$

Proof. Apply (C_2) and the following lemma. □

Lemma 2. *If* $\mathbf{A}$ *is a square matrix, then*

$$\operatorname{spr}\mathbf{A} \leqslant \max\left\{\operatorname{spr}\frac{z\,\mathbf{A} + \bar{z}\mathbf{A}^*}{2} \mid z \in \mathbf{C}, |z| = 1\right\}.$$

Proof of Lemma 2. Let $\lambda_j, \lambda_k \in \operatorname{spec}\mathbf{A}$ satisfy $|\lambda_j - \lambda_k| = \operatorname{spr}\mathbf{A}$. By Proposition 5, there exists a complex number z with $|z| = 1$, such that $(\lambda_j - \lambda_k)z = |\lambda_j - \lambda_k|$.
Now

$$\operatorname{spr}\mathbf{A} = |\lambda_j - \lambda_k| = (\lambda_j - \lambda_k)z = \operatorname{re}((\lambda_j - \lambda_k)z) = \operatorname{re}\lambda_j z - \operatorname{re}\lambda_k z.$$

Since $\lambda_j z, \lambda_k z \in \operatorname{spec} z\mathbf{A}$, we have $\operatorname{re}\lambda_j z, \operatorname{re}\lambda_k z \in \operatorname{re}\operatorname{spec} z\mathbf{A}$, and so, by Proposition 2, $\operatorname{re}\lambda_j z, \operatorname{re}\lambda_k z \in \operatorname{re} F(z\mathbf{A})$. On the other hand, by Propositions 1 and 2,

$$\operatorname{re} F(z\mathbf{A}) = F\left(\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\right) = \operatorname{co}\left(\operatorname{spec}\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\right),$$

and thus

$$\operatorname{spr}\mathbf{A} = \operatorname{re}\lambda_j z - \operatorname{re}\lambda_k z \leqslant \operatorname{diam}\operatorname{co}\left(\operatorname{spec}\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\right) = \operatorname{spr}\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2},$$

which proves the lemma. □

4. Generalizing (C_3)

Theorem 3. *If* $\mathbf{A}$ *is a normal* $n \times n$ *matrix, then*

(i) $\operatorname{spr}\mathbf{A} \geqslant \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{x}^*\mathbf{A}^*\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}$ *and*

(ii) $$\begin{aligned}\operatorname{spr}\mathbf{A} &= \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \operatorname{re}\mathbf{x}^*\mathbf{y} = 0\right\}\\ &= \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n, \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}.\end{aligned}$$

Proof. To show (i), apply (C_3) and (C_2) with $z = 1$. To show (ii), let $\mathbf{u}, \mathbf{v} \in \mathbf{C}^n$ satisfy $\|\mathbf{u}\|^2 = \|\mathbf{v}\|^2 = 1/2$, $\operatorname{re}\mathbf{u}^*\mathbf{v} = 0$. Then

$$\|\mathbf{u} \pm \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \pm 2 \operatorname{re} \mathbf{u}^*\mathbf{v} = 1,$$

and so, by Theorem 1,

$$\operatorname{spr} \mathbf{A} \geqslant |(\mathbf{u}+\mathbf{v})^*\mathbf{A}(\mathbf{u}+\mathbf{v}) - (\mathbf{u}-\mathbf{v})^*\mathbf{A}(\mathbf{u}-\mathbf{v})| = 2|\mathbf{u}^*\mathbf{A}\mathbf{v} + \mathbf{v}^*\mathbf{A}\mathbf{u}|.$$

Therefore, if $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ satisfy $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ and $\operatorname{re} \mathbf{x}^*\mathbf{y} = 0$, we, choosing $\mathbf{u} = \mathbf{x}/\sqrt{2}$, $\mathbf{v} = \mathbf{y}/\sqrt{2}$, obtain

$$\operatorname{spr} \mathbf{A} \geqslant |\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}|,$$

and so

$$\begin{aligned}\operatorname{spr} \mathbf{A} &\geqslant \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n,\ \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \operatorname{re} \mathbf{x}^*\mathbf{y} = 0\right\} \\ &\geqslant \max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y} + \mathbf{y}^*\mathbf{A}\mathbf{x}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n,\ \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\} = s.\end{aligned}$$

It remains to show that $s \geqslant \operatorname{spr} \mathbf{A}$. By ($C_1$), there exist $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$, such that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$, $\mathbf{x}^*\mathbf{y} = 0$, and $|\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathbf{y}^*\mathbf{A}\mathbf{y}| = \operatorname{spr} \mathbf{A}$. Then $\mathbf{u} = (\mathbf{x}+\mathbf{y})/\sqrt{2}$ and $\mathbf{v} = (\mathbf{x}-\mathbf{y})/\sqrt{2}$ satisfy $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$, $\mathbf{u}^*\mathbf{v} = 0$, and so

$$s \geqslant |\mathbf{u}^*\mathbf{A}\mathbf{v} + \mathbf{v}^*\mathbf{A}\mathbf{u}| = |\mathbf{x}^*\mathbf{A}\mathbf{x} - \mathbf{y}^*\mathbf{A}\mathbf{y}| = \operatorname{spr} \mathbf{A}. \qquad \square$$

Corollary 3. *If* $\mathbf{A}$ *is a Hermitian* $n \times n$ *matrix, then*

$$\begin{aligned}\operatorname{spr} \mathbf{A} &= 2 \max\left\{|\operatorname{re} \mathbf{x}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n,\ \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \operatorname{re} \mathbf{x}^*\mathbf{y} = 0\right\} \\ &= 2 \max\left\{|\operatorname{re} \mathbf{x}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n,\ \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}.\end{aligned}$$

5. Improving and generalizing (C_4)

Theorem 4. *If* $\mathbf{A}$ *is a Hermitian* $n \times n$ *matrix, then*

$$\operatorname{spr} \mathbf{A} = 2 \max\left\{(\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2)^{1/2} | \mathbf{x} \in \mathbf{C}^n,\ \|\mathbf{x}\| = 1\right\}.$$

Proof. The proof of (C_4) for real symmetric matrices [18, Theorem 1] works also for Hermitian matrices. Equality is attained for $\mathbf{x} = (\mathbf{u}_1 - \mathbf{u}_n)/\sqrt{2}$, since

$$\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2 = \tfrac{1}{2}(\lambda_1^2 + \lambda_n^2) - \tfrac{1}{4}(\lambda_1 + \lambda_n)^2 = \tfrac{1}{4}(\lambda_1 - \lambda_n)^2. \qquad \square$$

Corollary 4. *If* $\mathbf{A}$ *is a normal* $n \times n$ *matrix, then*

$$\operatorname{spr} \mathbf{A} = \sqrt{2} \max\left\{\left(|\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2| + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x} - |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2\right)^{1/2} \mid \mathbf{x} \in \mathbf{C}^n,\ \|\mathbf{x}\| = 1\right\}.$$

Proof. Take $\mathbf{x} \in \mathbf{C}^n$ with $\|\mathbf{x}\| = 1$ and $z \in \mathbf{C}$ with $|z| = 1$. Then

$$\mathbf{x}^* \left(\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\right)^2 \mathbf{x} = \frac{1}{4}\mathbf{x}^*\big(z^2\mathbf{A}^2 + \bar{z}^2(\mathbf{A}^*)^2 + \mathbf{A}^*\mathbf{A} + \mathbf{A}\mathbf{A}^*\big)\mathbf{x}$$
$$= \frac{1}{2}\big(\mathrm{re}(z^2\mathbf{x}^*\mathbf{A}^2\mathbf{x}) + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x}\big)$$

and

$$\left(\mathbf{x}^*\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\mathbf{x}\right)^2 = \frac{1}{4}\big(z\mathbf{x}^*\mathbf{A}\mathbf{x} + \bar{z}\mathbf{x}^*\mathbf{A}^*\mathbf{x}\big)^2 = \frac{1}{2}\big(\mathrm{re}(z\mathbf{x}^*\mathbf{A}\mathbf{x})^2 + |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2\big).$$

Hence, by Theorems 2 and 4,

$$\left(\frac{\operatorname{spr}\mathbf{A}}{2}\right)^2 \geqslant \left(\frac{1}{2}\operatorname{spr}\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\right)^2 \geqslant \mathbf{x}^*\left(\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\right)^2\mathbf{x} - \left(\mathbf{x}^*\frac{z\mathbf{A} + \bar{z}\mathbf{A}^*}{2}\mathbf{x}\right)^2$$
$$= \frac{1}{2}\big(\mathrm{re}(z^2\mathbf{x}^*\mathbf{A}^2\mathbf{x}) + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x}\big) - \frac{1}{2}\big(\mathrm{re}\,(z\mathbf{x}^*\mathbf{A}\mathbf{x})^2 + |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2\big)$$
$$= \frac{1}{2}\left\{\mathrm{re}\big[z^2\big(\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2\big)\big] + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x} - |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2\right\}.$$

Fix $\mathbf{x}$, and let z^2 be the z obtained applying Proposition 5 to $w = \mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2$. Then, by the above,

$$(*) \qquad \left(\frac{\operatorname{spr}\mathbf{A}}{2}\right)^2 \geqslant \frac{1}{2}\big(|\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2| + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x} - |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2\big).$$

It remains to show that equality is attained. Let $\operatorname{spr}\mathbf{A} = |\lambda_j - \lambda_k|$. Then $\mathbf{x} = (\mathbf{u}_j - \mathbf{u}_k)/\sqrt{2}$ satisfies

$$\big|\mathbf{x}^*\mathbf{A}^2\mathbf{x} - (\mathbf{x}^*\mathbf{A}\mathbf{x})^2\big| + \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x} - |\mathbf{x}^*\mathbf{A}\mathbf{x}|^2$$
$$= \big|\tfrac{1}{2}(\lambda_j^2 + \lambda_k^2) - \tfrac{1}{4}(\lambda_j + \lambda_k)^2\big| + \tfrac{1}{2}(|\lambda_j|^2 + |\lambda_k|^2) - \tfrac{1}{4}|\lambda_j + \lambda_k|^2$$
$$= \tfrac{1}{2}|\lambda_j - \lambda_k|^2,$$

which implies equality in $(*)$. □

6. An application of Theorem 1

Different choices of $\mathbf{x}$ and $\mathbf{y}$ in Theorem 1 give bounds for the spread of a normal matrix.

Johnson et al. [12] had (C_1) available. Assuming $\mathbf{A}$ real and normal, choosing $\mathbf{x} = \mathbf{e}/\sqrt{n}$, $\mathbf{y} = (\mathbf{e}_p - \mathbf{e}_q)/\sqrt{2}$ $(p \neq q)$, and doing some computation they obtained [12, Theorem 2.1]

$$\operatorname{spr} \mathbf{A} \geqslant \frac{1}{n-1}\left|\sum_j \sum_{k \neq j} a_{jk}\right|.$$

We show that reality of $\mathbf{A}$ is not needed.

Lemma 5. *If $w, z_1, \ldots, z_m$ are complex numbers, then*

$$\left|w - \frac{z_1 + \cdots + z_m}{m}\right| \leqslant \max_j |w - z_j|.$$

Proof. Trivial. □

Theorem 5. *If $\mathbf{A}$ is a normal $n \times n$ matrix, then*

$$\begin{aligned}\operatorname{spr} \mathbf{A} &\geqslant \max_{p \neq q}\left|\frac{1}{n}\sum_j \sum_k a_{jk} - \frac{a_{pp} + a_{qq} - a_{pq} - a_{qp}}{2}\right| \\ &\geqslant \frac{1}{n-1}\left|\sum_j \sum_{k \neq j} a_{jk}\right|.\end{aligned}$$

Proof. Applying Theorem 1 where $\mathbf{x} = \mathbf{e}/\sqrt{n}$, $\mathbf{y} = (\mathbf{e}_p - \mathbf{e}_q)/\sqrt{2}$ $(p \neq q)$ and taking the maximum over p, q, we have

$$\operatorname{spr} \mathbf{A} \geqslant \max_{p,q}\left|\frac{1}{n}\sum_j \sum_k a_{jk} - \frac{a_{pp} + a_{qq} - a_{pq} - a_{qp}}{2}\right|.$$

We underestimate the right-hand side. Applying Lemma 5 where $w = \frac{1}{n}\sum_j \sum_k a_{jk}$ and the z's are the $n(n-1)$ numbers

$$z_{pq} = \frac{a_{pp} + a_{qq} - a_{pq} - a_{qp}}{2} \quad (p \neq q),$$

we obtain

$$\begin{aligned}&\max_{p \neq q}\left|\frac{1}{n}\sum_j \sum_k a_{jk} - \frac{a_{pp} + a_{qq} - a_{pq} - a_{qp}}{2}\right| \\ &\quad\geqslant \left|\frac{1}{n}\sum_j \sum_k a_{jk} - \frac{1}{n(n-1)}\sum_p \sum_{q \neq p} \frac{a_{pp} + a_{qq} - a_{pq} - a_{qp}}{2}\right| \\ &\quad= \frac{1}{n-1}\left|\sum_j \sum_{k \neq j} a_{jk}\right|,\end{aligned}$$

which completes the proof. □

By Proposition 4, this theorem implies

Corollary 5. *If* $\mathbf{A}$ *is a normal* $n \times n$ *matrix, then*

$$\operatorname{spr} \mathbf{A} \geqslant \max_{I} \frac{1}{r-1} \left| \sum_{j \in I} \sum_{k \in I, k \neq j} a_{jk} \right|$$

where I *goes through the subsets of* $[n]$ *with* $r = |I| \geqslant 2$.

7. Improving Theorem 5

If $\mathbf{A}$ is symmetric and nonnegative, then the lower bound

$$\lambda_1 \geqslant \frac{\mathbf{e}^*\mathbf{A}\mathbf{e}}{n} = \frac{\operatorname{su} \mathbf{A}}{n}$$

is often amazingly good [14]. If $\mathbf{A}$ is also nonzero, the bound

$$\lambda_1 \geqslant \frac{(\mathbf{A}\mathbf{e})^*\mathbf{A}(\mathbf{A}\mathbf{e})}{(\mathbf{A}\mathbf{e})^*(\mathbf{A}\mathbf{e})} = \frac{\operatorname{su} \mathbf{A}^3}{\operatorname{su} \mathbf{A}^2}$$

is still better in most cases, and if $n \leqslant 3$ in all cases [13]. This motivates us to try to improve Theorem 5.

Let $\mathbf{A}$ be a real and symmetric matrix whose row sum vector $\mathbf{r} \neq \mathbf{0}$. We choose in Theorem 1

$$\mathbf{x} = \frac{\mathbf{A}\mathbf{e}}{\sqrt{(\mathbf{A}\mathbf{e})^*\mathbf{A}\mathbf{e}}} = \frac{\mathbf{r}}{\sqrt{\mathbf{r}^*\mathbf{r}}}$$

and $\mathbf{y} = \mathbf{e}_p \cos\theta + \mathbf{e}_q \sin\theta$ $(p \neq q)$. Then the minimum of

$$f(\theta) = \mathbf{y}^*\mathbf{A}\mathbf{y} = a_{pp}\cos^2\theta + 2a_{pq}\sin\theta\cos\theta + a_{qq}\sin^2\theta = \mathbf{z}^{\mathrm{T}}\mathbf{A}_{pq}\mathbf{z},$$

where $\mathbf{z}^{\mathrm{T}} = (\cos\theta \; \sin\theta)$ and

$$\mathbf{A}_{pq} = \begin{pmatrix} a_{pp} & a_{pq} \\ a_{pq} & a_{qq} \end{pmatrix},$$

is

$$\frac{a_{pp} + a_{qq} - \sqrt{(a_{pp} - a_{qq})^2 + 4a_{pq}^2}}{2},$$

the smaller eigenvalue of $\mathbf{A}_{pq}$. Thus we may expect that the bound

$$\operatorname{spr} \mathbf{A} \geqslant \frac{\operatorname{su} \mathbf{A}^3}{\operatorname{su} \mathbf{A}^2} - \min_{p \neq q} \frac{a_{pp} + a_{qq} - \sqrt{(a_{pp} - a_{qq})^2 + 4a_{pq}^2}}{2}$$

is often good if $\mathbf{A}$ is also nonnegative.

More generally, we have

Theorem 6. *If* $\mathbf{A}$ *is a Hermitian* $n \times n$ *matrix whose row sum vector is nonzero*, *then*

$$\operatorname{spr}\mathbf{A} \geqslant \max_{p\neq q}\left|\frac{\operatorname{su}\mathbf{A}^3}{\operatorname{su}\mathbf{A}^2} - \frac{a_{pp}+a_{qq} \pm \sqrt{(a_{pp}-a_{qq})^2+4|a_{pq}|^2}}{2}\right|$$
$$\geqslant \left|\frac{\operatorname{su}\mathbf{A}^3}{\operatorname{su}\mathbf{A}^2} - \frac{\operatorname{tr}\mathbf{A}}{n}\right|.$$

Proof. Putting $\mathbf{x}$ as above and $\mathbf{y} = \mathbf{e}_p s + \mathbf{e}_q t$ $[\mathbf{y} = \mathbf{e}_p s' + \mathbf{e}_q t']$ where $(s, t)^{\mathrm{T}}$ $[(s', t')^{\mathrm{T}}]$ is a unit eigenvector corresponding to the larger [smaller] eigenvalue of $\mathbf{A}_{pq}$ $(p \neq q)$, applying Theorem 1, and maximizing over p, q, we obtain the first inequality. The second inequality follows by applying Lemma 5 where

$$w = \frac{\operatorname{su}\mathbf{A}^3}{\operatorname{su}\mathbf{A}^2}$$

and the z's are the $2n(n-1)$ numbers

$$z_{pq} = \frac{a_{pp}+a_{qq} \pm \sqrt{(a_{pp}-a_{qq})^2+4|a_{pq}|^2}}{2} \quad (p \neq q).$$

Putting $\mathbf{x} = \mathbf{e}_p s + \mathbf{e}_q t$, $\mathbf{y} = \mathbf{e}_p s' + \mathbf{e}_q t'$ we obtain similarly

$$(*) \quad \operatorname{spr}\mathbf{A} \geqslant \max_{p\neq q}\sqrt{(a_{pp}-a_{qq})^2+4|a_{pq}|^2},$$

proved by Mirsky [16, Theorem 2] using Proposition 4. □

Still more generally, we can similarly prove

Theorem 7. *If* $\mathbf{A}$ *is a normal* $n \times n$ *matrix whose row sum vector is nonzero*, *then*

$$\operatorname{spr}\mathbf{A} \geqslant \max_{p\neq q}\left|\frac{\operatorname{su}\mathbf{A}^*\mathbf{A}^2}{\operatorname{su}\mathbf{A}^*\mathbf{A}} - \frac{a_{pp}+a_{qq} \pm \sqrt{(a_{pp}-a_{qq})^2+4a_{pq}a_{qp}}}{2}\right|$$
$$\geqslant \left|\frac{\operatorname{su}\mathbf{A}^*\mathbf{A}^2}{\operatorname{su}\mathbf{A}^*\mathbf{A}} - \frac{\operatorname{tr}\mathbf{A}}{n}\right|.$$

Studying $\operatorname{spr}(z\mathbf{A}) = \operatorname{spr}\mathbf{A}$ $(|z| = 1)$ as in the proof of $(*)$ and choosing z suitably (Proposition 5), it can be shown for normal matrices that

$$\operatorname{spr}\mathbf{A} \geqslant \max_{p\neq q}\sqrt{|(a_{pp}-a_{qq})^2+4a_{pq}a_{qp}|}.$$

However, applying Theorem 2 to (∗) with a suitable z gives a better bound

$$\operatorname{spr}\mathbf{A} \geqslant \frac{1}{\sqrt{2}} \max_{p\neq q} \big[|a_{pp}-a_{qq}|^2 + |(a_{pp}-a_{qq})^2 + 4a_{pq}a_{qp}| + 2|a_{pq}|^2 + 2|a_{qp}|^2\big]^{1/2},$$

proved differently by Mirsky [16, Theorem 4].

8. Another application of Theorem 1 and an application of Theorem 3

Johnson et al. [12, Theorem 2.2] proved for normal matrices that

$$\operatorname{spr}\mathbf{A} \geqslant \max_{J,K} \left|\frac{1}{r}\operatorname{su}\mathbf{A}_J - \frac{1}{s}\operatorname{su}\mathbf{A}_K\right|$$

and for Hermitian matrices that

$$\operatorname{spr}\mathbf{A} \geqslant \max_{J,K} \frac{2}{\sqrt{rs}}\,|\operatorname{su}\mathbf{A}_{JK}|\,,$$

where J and K go through all disjoint nonempty subsets of $[n]$ and $r = |J|$, $s = |K|$. We extend the first inequality to nondisjoint subsets, and the second inequality to normal matrices.

Theorem 8. *If* $\mathbf{A}$ *is a normal* $n \times n$ *matrix, then*

(i) $\operatorname{spr}\mathbf{A} \geqslant \max_{J,K} \left|\frac{1}{r}\operatorname{su}\mathbf{A}_J - \frac{1}{s}\operatorname{su}\mathbf{A}_K\right|,$

where J *and* K *go through all nonempty subsets of* $[n]$ *and* $r = |J|$, $s = |K|$. *Furthermore,*

(ii) $\operatorname{spr}\mathbf{A} \geqslant \max_{J,K} \frac{1}{\sqrt{rs}}\big|\operatorname{su}\mathbf{A}_{JK} + \operatorname{su}\bar{\mathbf{A}}_{JK}\big|$ *and*

(iii) $\operatorname{spr}\mathbf{A} \geqslant \max_{J,K} \frac{1}{\sqrt{rs}}\big|\operatorname{su}\mathbf{A}_{JK} + \operatorname{su}\mathbf{A}_{KJ}\big|,$

where J *and* K *are as above but disjoint.*

Proof. Put $\mathbf{x} = \mathbf{e}_J/\sqrt{r}$, $\mathbf{y} = \mathbf{e}_K/\sqrt{s}$, and apply Theorem 1 and Theorem 3, respectively. □

To see that including nondisjoint J, K in (i) is indeed an improvement, consider

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 0 \end{pmatrix}.$$

Then $\max_{J,K}\left|\frac{1}{r}\operatorname{su}\mathbf{A}_J - \frac{1}{s}\operatorname{su}\mathbf{A}_K\right|$ over disjoint subsets is 3 ($J = \{1, 2, 3\}$, $K = \{4\}$), while over all subsets it is $3\frac{1}{2}$ ($J = \{1, 2, 3\}$, $K = \{3, 4\}$).

9. Applications of Theorem 4 and Corollary 4

Different choices of $\mathbf{x}$ in Theorem 4 and Corollary 4 give bounds for the spread of a Hermitian matrix and of a normal matrix, respectively.

Let $\mathbf{A}$ be a normal $n \times n$ matrix with real row sums $r_1, \ldots, r_n$. Johnson et al. [12, Theorem 2.3] proved that

$$\operatorname{spr}\mathbf{A} \geqslant \sqrt{3}\left[\frac{1}{n}\sum_j r_j^2 - \left(\frac{1}{n}\sum_j r_j\right)^2\right]^{1/2}.$$

In (iii) below we extend this to complex row sums.

Theorem 9. *Let* $\mathbf{A}$ *be a normal* $n \times n$ *matrix. Then*

(i) $$\operatorname{spr}\mathbf{A} \geqslant \sqrt{2}\max_j\left(\left|\sum_{k\neq j} a_{jk}a_{kj}\right| + \sum_{k\neq j}|a_{jk}|^2\right)^{1/2}.$$

If $r_1, \ldots, r_n$ *and* $c_1, \ldots, c_n$ *are respectively the row and column sums of* $\mathbf{A}$, *then*

(ii) $$\operatorname{spr}\mathbf{A} \geqslant \sqrt{2}\left[\left|\frac{\operatorname{su}\mathbf{A}^2}{n} - \left(\frac{\operatorname{su}\mathbf{A}}{n}\right)^2\right| + \frac{\operatorname{su}\mathbf{A}^*\mathbf{A}}{n} - \left|\frac{\operatorname{su}\mathbf{A}}{n}\right|^2\right]^{1/2}$$
$$= \sqrt{2}\left[\left|\frac{1}{n}\sum_j r_jc_j - \left(\frac{1}{n}\sum_j r_j\right)^2\right| + \frac{1}{n}\sum_j |r_j|^2 - \left|\frac{1}{n}\sum_j r_j\right|^2\right]^{1/2} \textit{ and}$$

(iii) $$\operatorname{spr}\mathbf{A} \geqslant \sqrt{3}\left(\frac{\operatorname{su}\mathbf{A}^*\mathbf{A}}{n} - \left|\frac{\operatorname{su}\mathbf{A}}{n}\right|^2\right)^{1/2} = \sqrt{3}\left(\frac{1}{n}\sum_j |r_j|^2 - \left|\frac{1}{n}\sum_j r_j\right|^2\right)^{1/2}.$$

Proof. (i) Apply Corollary 4 with $\mathbf{x} = \mathbf{e}_j$ and maximize over the j's. (ii) Apply Corollary 4 with $\mathbf{x} = \mathbf{e}/\sqrt{n}$. (iii) Recall

$$(\mathrm{C}_3') \quad \operatorname{spr}\mathbf{A} \geqslant \sqrt{3}\max\left\{|\mathbf{x}^*\mathbf{A}\mathbf{y}| \mid \mathbf{x}, \mathbf{y} \in \mathbf{C}^n,\ \|\mathbf{x}\| = \|\mathbf{y}\| = 1, \mathbf{x}^*\mathbf{y} = 0\right\}.$$

If

$$v = \frac{1}{n}\sum_j |r_j|^2 - \left|\frac{1}{n}\sum_j r_j\right|^2 = 0,$$

then the claim is trivially true. Otherwise ($v > 0$), put $\mathbf{y} = \mathbf{e}/\sqrt{n}$ and $\mathbf{x} = (x_j)$ where

$$x_j = \frac{r_j - \frac{1}{n}\sum_k r_k}{\sqrt{nv}}.$$

(It is easy to see that indeed $\|\mathbf{x}\| = 1$.) $\square$

Corollary 9. *If* $\mathbf{A}$ *is a Hermitian* $n \times n$ *matrix with row sums* $r_1, \dots, r_n$, *then*

(i) $\operatorname{spr} \mathbf{A} \geqslant 2 \max_j \left(\sum_{k \neq j} |a_{jk}|^2 \right)^{1/2}$ *and*

(ii) $\operatorname{spr} \mathbf{A} \geqslant 2 \left[\frac{\operatorname{su} \mathbf{A}^2}{n} - \left(\frac{\operatorname{su} \mathbf{A}}{n} \right)^2 \right]^{1/2} = 2 \left[\frac{1}{n} \sum_j |r_j|^2 - \left(\frac{1}{n} \sum_j r_j \right)^2 \right]^{1/2}.$

Johnson et al. [12, Theorem 2.3] proved (ii) assuming the row sums real.

10. Two further applications of Theorem 3

We present more bounds using row and column sums.

Theorem 10. *Let* $\mathbf{A}$ *be a normal* $n \times n$ *matrix with row sums* $r_1, \dots, r_n$ *and column sums* $c_1, \dots, c_n$, *and denote*

$$v = \frac{\operatorname{su} \mathbf{A}^*\mathbf{A}}{n} - \left| \frac{\operatorname{su} \mathbf{A}}{n} \right|^2 = \frac{1}{n} \sum_j |r_j|^2 - \left| \frac{1}{n} \sum_j r_j \right|^2,$$

$$w = \frac{\operatorname{su} \mathbf{A}^2}{n} - \left(\frac{\operatorname{su} \mathbf{A}}{n} \right)^2 = \frac{1}{n} \sum_j r_j c_j - \left(\frac{1}{n} \sum_j r_j \right)^2.$$

If $v > 0$, *then*

$$\operatorname{spr} \mathbf{A} \geqslant \left| \sqrt{v} + \frac{w}{\sqrt{v}} \right|.$$

Sketch of Proof. Choose $\mathbf{x}$ and $\mathbf{y}$ as in the proof of Theorem 9(iii) and apply Theorem 3(ii). $\square$

For Hermitian matrices, Theorem 10 implies Corollary 9(ii) again.

Johnson et al. [12, Theorem 2.4] proved also the following theorem assuming the row sums real. (In [12, (2.14)], for normal $\mathbf{A}$, their $K = \frac{3}{2}$, though not wrong, should be replaced with $K = \sqrt{3}$.)

Theorem 11. *Let* $\mathbf{A}$ *be a normal* $n \times n$ *matrix with row sums* $r_1, \dots, r_n$, *let* $\sigma = (j_1, \dots, j_n)$ *be a permutation of* $(1, \dots, n)$, *and denote*

$$m = \left\lfloor \frac{n+1}{2} \right\rfloor.$$

Define $\boldsymbol{\rho} = (\rho_t)$ *by* $\rho_t = r_{j_t} - r_{j_{n-t+1}}$, $t = 1, \dots, m$. *Then*

(i) $\operatorname{spr}\mathbf{A} \geqslant \max_\sigma \left(\frac{3}{2n} \sum_{t=1}^{m} \left| r_{j_t} - r_{j_{n-t+1}} \right|^2 \right)^{1/2} = \sqrt{\frac{3}{2n}} \max_\sigma \|\boldsymbol{\rho}\|$ *and*

(ii) $\operatorname{spr}\mathbf{A} \geqslant \max_\sigma \frac{\sqrt{3}}{n} \sum_{t=1}^{m} \left| r_{j_t} - r_{j_{n-t+1}} \right| = \frac{\sqrt{3}}{n} \max_\sigma \|\boldsymbol{\rho}\|_1.$

If $\mathbf{A}$ *is Hermitian, then*

(iii) $\operatorname{spr}\mathbf{A} \geqslant \max_\sigma \left(\frac{2}{n} \sum_{t=1}^{m} \left| r_{j_t} - r_{j_{n-t+1}} \right|^2 \right)^{1/2} = \sqrt{\frac{2}{n}} \max_\sigma \|\boldsymbol{\rho}\|$ *and*

(iv) $\operatorname{spr}\mathbf{A} \geqslant \max_\sigma \frac{2}{n} \sum_{t=1}^{m} \left| r_{j_t} - r_{j_{n-t+1}} \right| = \frac{2}{n} \max_\sigma \|\boldsymbol{\rho}\|_1.$

Sketch of Proof. Exclude the trivial case of equal row sums. To show (i) and (iii), define $\boldsymbol{\rho}' = (\rho_t)$ with $\rho_t = r_{j_t} - r_{j_{n-t+1}}$, $t = 1, \dots, n$, choose $\mathbf{x} = \boldsymbol{\rho}'/\|\boldsymbol{\rho}'\|$, $\mathbf{y} = \mathbf{e}/\sqrt{n}$, and apply ($C_3'$) and ($C_3$). To show (ii) and (iv), proceed similarly, but choose $\mathbf{x}$ as follows: For $1 \leqslant t \leqslant m$, take z_t satisfying $|z_t| = 1$ and $\bar{z}_t \rho_t = |\rho_t|$. Put $\mathbf{x} = (z_1, \dots, z_m, -z_m, \dots, -z_1)^{\mathrm{T}}/\sqrt{n}$ if n is even, and insert an arbitrary z with $|z| = 1$ in the middle if n is odd. The bound (iv) follows also from (iii) by the Cauchy–Schwarz inequality, and so (iii) is better than (iv). □

Theorem 12. *Let* $\mathbf{A}$ *be a normal* $n \times n$ *matrix with row sums* $r_1, \dots, r_n$ *and column sums* $c_1, \dots, c_n$. *Define* σ, m, *and* $\boldsymbol{\rho}$ *as in Theorem* 11, *and* $\boldsymbol{\gamma} = (\gamma_t)$ *by* $\gamma_t = c_{j_t} - c_{j_{n-t+1}}$, $t = 1, \dots, m$. *Then*

(i) $\operatorname{spr}\mathbf{A} \geqslant \frac{1}{\sqrt{n}} \max_\sigma \left| \frac{\|\boldsymbol{\rho}\|}{\sqrt{2}} + \frac{\boldsymbol{\rho}^{\mathrm{T}}\boldsymbol{\gamma}}{\|\boldsymbol{\rho}\|} \right|$ (*assuming* $\boldsymbol{\rho} \neq \mathbf{0}$) *and*

(ii) $\operatorname{spr}\mathbf{A} \geqslant \frac{1}{n} \max_\sigma \|\boldsymbol{\rho} + \bar{\boldsymbol{\gamma}}\|_1$.

Sketch of Proof. To show (i), take $\mathbf{x}$ and $\mathbf{y}$ as in the proof of Theorem 11(i) and (iii), and apply Theorem 3(i) (or 3(ii)). To show (ii), take $\mathbf{x}$ and $\mathbf{y}$ as in the proof of Theorem 11(ii) and (iv) but take z_t satisfying $|z_t| = 1$ and $\bar{z}_t(\rho_t + \bar{\gamma}_t) = |\rho_t + \bar{\gamma}_t|$, and apply Theorem 3(i). □

11. Some other bounds

Theorem 13. *If* **A** *is Hermitian, then*

(i) $\operatorname{spr}\mathbf{A} \geqslant \max_{j,k}\left[(a_{jj}-a_{kk})^2 + 2\sum_{p\neq j}|a_{jp}|^2 + 2\sum_{p\neq k}|a_{kp}|^2\right]^{1/2}$ (Barnes and Hoffman, [1, (2.6)]),

(ii) $\operatorname{spr}\mathbf{A} \geqslant \frac{1}{\sqrt{n}}\max_{j,k}\left(|a_{jj}-a_{kk}| + \sum_{p\neq j}|a_{jp}| + \sum_{p\neq k}|a_{kp}|\right)$ (Scott [17, Theorem1], see also [1, (3.3)]).

The bound (i) improves Corollary 9(i), since

$$\max_{j,k}\left[(a_{jj}-a_{kk})^2 + 2\sum_{p\neq j}|a_{jp}|^2 + 2\sum_{p\neq k}|a_{kp}|^2\right] \geqslant 4\max_{j}\sum_{p\neq j}|a_{jp}|^2.$$

Jiang and Zhan [10, Theorems 4 and 5] improved (i), but to extend the improved bounds to normal matrices seems too complicated. To extend (ii) to normal matrices seems also too complicated. We extend (i) to normal matrices.

Theorem 14. *If* **A** *is a normal matrix, then*

$$\begin{aligned}\operatorname{spr}\mathbf{A} \geqslant \frac{1}{\sqrt{2}}\max_{j,k}\Bigg[&\left|(a_{jj}-a_{kk})^2 + 2\sum_{p\neq j}a_{jp}a_{pj} + 2\sum_{p\neq k}a_{kp}a_{pk}\right|\\ &+|a_{jj}-a_{kk}|^2 + \sum_{p\neq j}|a_{jp}|^2 + \sum_{p\neq j}|a_{pj}|^2 + \sum_{p\neq k}|a_{kp}|^2\\ &+\sum_{p\neq k}|a_{pk}|^2\Bigg]^{1/2}.\end{aligned}$$

Proof. Let $|z| = 1$ and $1 \leqslant j, k \leqslant n$. Applying Theorem 13(i) to $(z\mathbf{A} + \bar{z}\mathbf{A}^*)/2$, using ($\mathrm{C}_2$), and choosing z suitably (cf. the proof of Corollary 4), we have

$$\begin{aligned}(\operatorname{spr}\mathbf{A})^2 \geqslant{}& \left[\frac{z(a_{jj}-a_{kk}) + \bar{z}(\bar{a}_{jj}-\bar{a}_{kk})}{2}\right]^2 + 2\sum_{p\neq j}\left|\frac{za_{jp}+\bar{z}\bar{a}_{pj}}{2}\right|^2\\ &+2\sum_{p\neq k}\left|\frac{za_{kp}+\bar{z}\bar{a}_{pk}}{2}\right|^2\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\mathrm{re}\left[z^2\left(2\sum_{p\neq j} a_{jp}a_{pj} + 2\sum_{p\neq k} a_{kp}a_{pk} + (a_{jj}-a_{kk})^2\right)\right] \\
&\quad + \frac{1}{2}\left(\sum_{p\neq j}|a_{jp}|^2 + \sum_{p\neq j}|a_{pj}|^2 + \sum_{p\neq k}|a_{kp}|^2 + \sum_{p\neq k}|a_{pk}|^2 + |a_{jj}-a_{kk}|^2\right) \\
&= \frac{1}{2}\left|2\sum_{p\neq j} a_{jp}a_{pj} + 2\sum_{p\neq k} a_{kp}a_{pk} + (a_{jj}-a_{kk})^2\right| \\
&\quad + \frac{1}{2}\left(\sum_{p\neq j}|a_{jp}|^2 + \sum_{p\neq j}|a_{pj}|^2 + \sum_{p\neq k}|a_{kp}|^2 + \sum_{p\neq k}|a_{pk}|^2 + |a_{jj}-a_{kk}|^2\right).
\end{aligned}
$$

□

12. Bounds involving traces

Studying $\mathbf{u}_1, \ldots, \mathbf{u}_n$ instead of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ shows that our bounds involving element sums have analogies involving traces. The following bounds are easily obtained.

Theorem 15 (cf. Theorem 6). *If* $\mathbf{A}$ *is a nonzero Hermitian* $n \times n$ *matrix, then*

$$\operatorname{spr}\mathbf{A} \geqslant \left|\frac{\operatorname{tr}\mathbf{A}^3}{\operatorname{tr}\mathbf{A}^2} - \frac{\operatorname{tr}\mathbf{A}}{n}\right|.$$

Theorem 16 (cf. Theorem 9). *If* $\mathbf{A}$ *is a normal* $n \times n$ *matrix, then*

$$\operatorname{spr}\mathbf{A} \geqslant \sqrt{3}\left(\frac{\operatorname{tr}\mathbf{A}^*\mathbf{A}}{n} - \left|\frac{\operatorname{tr}\mathbf{A}}{n}\right|^2\right)^{1/2}.$$

For Hermitian matrices, this theorem can be improved.

Theorem 17 (Brauer and Mewborn [4], cf. Corollary 9). *If* $\mathbf{A}$ *is an* $n \times n$ *matrix with real eigenvalues, then*

$$\operatorname{spr}\mathbf{A} \geqslant 2\left[\frac{\operatorname{tr}\mathbf{A}^2}{n} - \left(\frac{\operatorname{tr}\mathbf{A}}{n}\right)^2\right]^{1/2}.$$

If n is odd, a further improvement exists.

Theorem 18 (Wolkowicz and Styan [20, (2.50)]). *If* $\mathbf{A}$ *is an* $n \times n$ *matrix with real eigenvalues and* n *is odd, then*

$$\operatorname{spr}\mathbf{A} \geqslant 2\left[\frac{n\operatorname{tr}\mathbf{A}^2 - (\operatorname{tr}\mathbf{A})^2}{n^2-1}\right]^{1/2}.$$

13. Remarks

All our bounds (restricting to fixed subsets in Corollary 5 and Theorem 8, and to a fixed permutation in Theorems 11 and 12) have complexity $\mathrm{O}(n^2)$. Note that $\operatorname{su}\mathbf{A}^2$, $\operatorname{su}\mathbf{A}^*\mathbf{A}$, $\operatorname{su}\mathbf{A}^*\mathbf{A}^2$, $\operatorname{tr}\mathbf{A}^2$, and $\operatorname{tr}\mathbf{A}^*\mathbf{A}$ can be found without computing these matrices.

Instead of normality, it is in fact enough to assume only that

$$(*) \quad F(\mathbf{A}) = \operatorname{co}(\operatorname{spec}\mathbf{A}).$$

This requires modification only in Corollary 4 where $\mathbf{A}^*\mathbf{A}$ must be replaced with $(\mathbf{A}^*\mathbf{A} + \mathbf{A}\mathbf{A}^*)/2$.

Normality implies $(*)$, see Proposition 2, but $(*)$ does not necessarily imply normality for $n \geqslant 5$, see [7, Problem 1.2.10]. Matrices satisfying $(*)$ are characterized by Johnson [11], see also [7, Theorem 1.6.8].

14. Experiments

To compare the bounds discussed in previous sections, we performed one hundred numerical experiments with random matrices of order 25, using the Matlab random generators rand and randn. We did not include bounds which are more complex than $\mathrm{O}(n^2)$. Therefore we omitted Corollary 5 and studied Theorem 8 with fixed subsets and Theorems 11 and 12 with a fixed permutation. The results below give means (m) and standard deviations (s) of relative errors

$$\frac{\operatorname{spr}\mathbf{A} - \text{bound}}{\operatorname{spr}\mathbf{A}}$$

of the best and second best bounds for each matrix type.

Type	Positive symmetric
Generator	rand
Best	Theorem 6 ($m = 0.1108$, $s = 0.0111$)
Second best	Theorem 5 ($m = 0.1205$, $s = 0.0110$)

Type	Real symmetric
Generator	randn
Best	Theorem 13(i) ($m = 0.3023$, $s = 0.0395$)
Second best	Corollary 9(i) ($m = 0.3041$, $s = 0.0412$)

Type	Hermitian; positive real and imaginary parts of the upper triangle
Generator	rand
Best	Theorem 6 ($m = 0.2351$, $s = 0.0162$)
Second best	Corollary 9(ii) ($m = 0.3218$, $s = 0.0206$)

Type	Hermitian
Generator	randn
Best	Theorem 13(i) ($m = 0.3385$, $s = 0.0333$)
Second best	Corollary 9(i) ($m = 0.3392$, $s = 0.0338$)

Type	Normal; positive real and imaginary parts of eigenvalues
Generator	rand (in generating eigenvalues)
Best	Theorem 14 ($m = 0.3604$, $s = 0.0460$)
Second best	Theorem 9(i) ($m = 0.3626$, $s = 0.0461$)

Type	Normal
Generator	randn (in generating eigenvalues)
Best	Theorem 14 ($m = 0.4092$, $s = 0.0457$)
Second best	Theorem 9(i) ($m = 0.4107$, $s = 0.0455$)

All the bounds on this top list except Corollary 9(ii) are produced by maximizing a bound over rows and (or) columns. It is no surprise that such bounds beat bounds obtained by a single formula. In the case of positive symmetric matrices, Theorems 6 and 5 give, in average, amazingly good bounds. This is apparently reminiscent of the fact that su $\mathbf{A}/n$ is often an amazingly good lower bound for the largest eigenvalue of such matrices (see Section 6). Also positivity of real and imaginary parts of the upper triangle in the Hermitian case and of eigenvalues in the normal case improves results.

Acknowledgments

We thank Professor George P.H. Styan for alerting to us several references. We also thank the referees for their thorough work.

References

[1] E.R. Barnes, A.J. Hoffman, Bounds for the spectrum of normal matrices, Linear Algebra Appl. 201 (1994) 79–90.
[2] P.R. Beesack, The spread of matrices and polynomials, Linear Algebra Appl. 31 (1980) 145–149.
[3] P. Bloomfield, G.S. Watson, The inefficiency of least squares, Biometrika 62 (1975) 121–128.
[4] A. Brauer, A.C. Mewborn, The greatest distance between two characteristic roots of a matrix, Duke Math. J. 26 (1959) 653–661.
[5] E. Deutsch, On the spread of matrices and polynomials, Linear Algebra Appl. 22 (1978) 49–55.
[6] R. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[7] R. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[8] S.T. Jensen, The Laguerre–Samuelson inequality with extensions and applications in statistics and matrix theory, Master's thesis, McGill University, Montreal, 1999.

[9] Z. Jia, An extension of Styan's inequality, Gongcheng Shuxue Xuebao 13 (1996) 122–126 (Chinese, English summary).
[10] E. Jiang, X. Zhan, Lower bounds for the spread of a Hermitian matrix, Linear Algebra Appl. 256 (1997) 153–163.
[11] C.R. Johnson, Normality and the numerical range, Linear Algebra Appl. 15 (1976) 89–94.
[12] C.R. Johnson, R. Kumar, H. Wolkowicz, Lower bounds for the spread of a matrix, Linear Algebra Appl. 71 (1985) 161–173.
[13] D. London, Two inequalities in nonnegative symmetric matrices, Pacific J. Math. 16 (1966) 515–536.
[14] J.K. Merikoski, On a lower bound for the Perron eigenvalue, BIT 19 (1979) 39–42.
[15] L. Mirsky, The spread of a matrix, Mathematika 3 (1956) 127–130.
[16] L. Mirsky, Inequalities for normal and Hermitian matrices, Duke Math. J. 24 (1957) 591–599.
[17] D.S. Scott, On the accuracy of the Gershgorin circle theorem for bounding the spread of a real symmetric matrix, Linear Algebra Appl. 65 (1985) 147–155.
[18] G.P.H. Styan, On some inequalities associated with ordinary least squares and the Kantorovich inequality, Acta Univ. Tamper. Ser. A 153 (1983) 158–166.
[19] B. Tu, On the spread of a matrix, Fudan Xuebao, Ziran Kexue Ban 23 (1984) 435–442 (Chinese).
[20] H. Wolkowicz, G.P.H. Styan, Bounds for eigenvalues using traces, Linear Algebra Appl. 29 (1980) 471–506.